IDOL: INVERSE-DYNAMICS-GUIDED FUTURE PREDICTION FOR END-TO-END AUTONOMOUS DRIVING

Anonymous authors
Paper under double-blind review

ABSTRACT

End-to-end autonomous driving has emerged as a compelling paradigm for learning planning directly from sensor observations, while recent world-model-based approaches further enrich this paradigm by enabling explicit reasoning about how the scene may evolve in the future. Yet future prediction alone does not guarantee better planning unless the predicted evolution can be converted into planning-relevant trajectory updates. Many current methods still forecast future scene states without explicitly decoding the motion implications hidden in state transitions. As a result, future reasoning often remains descriptively useful but only weakly coupled to executable motion generation. To address this limitation, we propose **IDOL**, an inverse-dynamics-guided future prediction framework for world-model-based end-to-end planning in latent BEV space, where inverse dynamics serves as the key bridge between future prediction and trajectory optimization. IDOL first predicts multiple future latent scene states with a BEV world model, then applies an inverse dynamics model to adjacent latent futures to decode transition-aware trajectory features and recover planning-relevant motion deltas that explain how the latent world evolves over time. These transition features refine the candidate-specific planning queries through spatial cross-attention and global modulation. A lightweight closed-loop refinement module feeds the refined queries back into candidate-conditioned latent rollouts, enabling an additional round of future-informed query refinement before trajectory decoding. By introducing inverse dynamics into latent future reasoning, IDOL tightens the coupling between world modeling and planning. Extensive experiments on the NAVSIM v1 and NAVSIM v2 benchmarks show that IDOL achieves state-of-the-art performance among comparable methods.

1 INTRODUCTION

Autonomous driving (Jiang et al., 2023; Chen et al., 2024b; Sun et al., 2025b; Li et al., 2024b; Liu et al., 2025b; Guo et al., 2026; Huang et al., 2026; Tang et al., 2025) has advanced significantly in recent years, driven by a paradigm shift from modular perception-prediction-planning pipelines to unified end-to-end learning frameworks. Recent advances have further expanded this paradigm toward future-aware, world-model-guided planning (Zhang et al., 2026; Xia et al., 2025), scalable sim-real training (Tian et al., 2025), and video-centric pretraining (Strong et al., 2026), leading to markedly stronger robustness and planning performance on realistic benchmarks such as nuScenes (Caesar et al., 2020) and NAVSIM (Dauner et al., 2024). At the same time, multi-modal and generative formulations (Xu et al., 2025; Liu et al., 2025a; Li et al., 2025c; Zhao et al., 2025) have shown that planning quality can benefit substantially from moving beyond deterministic single-pass regression.More recently, VLA-based systems (Li et al., 2026; Jiang et al., 2025a; Luo et al., 2025) extended the scope of future-aware decision making by injecting language-grounded reasoning, action abstraction, and reflective post-training into driving systems.

For trajectory planning, traditional end-to-end autonomous driving methods (Tang et al., 2026; Wang et al., 2026b) typically rely on the current state, which constitutes a form of passive planning, as illustrated in Figure 1(a). Subsequently, many approaches (Jiang et al., 2025c; Kim et al., 2026) have introduced world models into end-to-end planning, adopting the predict-decide paradigm to make decision-making more foresighted, as illustrated in Figure 1(b).

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

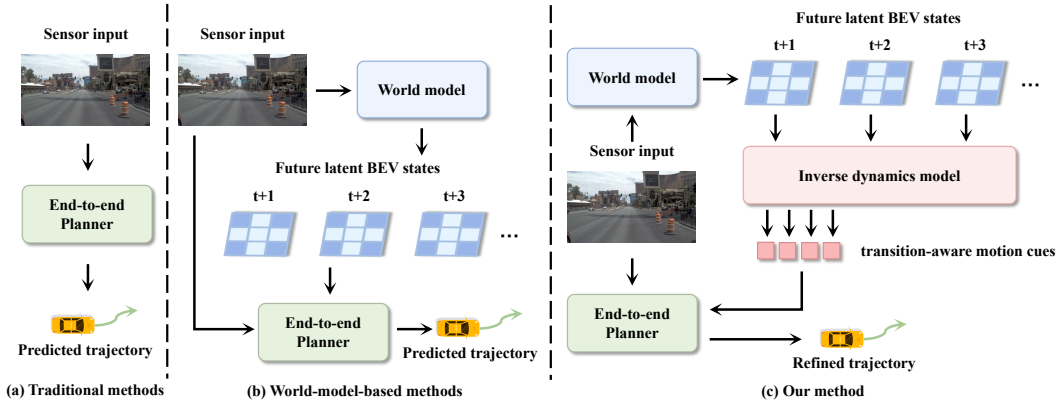


Figure 1: Comparison of three planning paradigms. (a) Traditional methods directly map sensor inputs to a predicted trajectory. (b) World-model-based methods additionally predict future latent BEV states to support planning, but these predicted futures are only loosely connected to motion generation. (c) In contrast, IDOL further applies an inverse dynamics model to adjacent future latent BEV states to extract transition-aware motion cues for trajectory refinement, making future prediction more actionable for planning.

Yet a fundamental bottleneck remains: even with future prediction, current methods (Li et al., 2025e; Zhang et al., 2025a) still lack a clear and interpretable bridge between forecasting scene evolution and using it to produce dynamically meaningful trajectory updates. Predicted futures may therefore be informative but insufficiently actionable: the planner can anticipate what may happen next, yet lacks a principled mechanism to revise the current trajectory so that it remains executable, dynamically coherent, and robust over long horizons. This limitation is especially critical in safety-sensitive driving, where future reasoning must ultimately support motion generation rather than passive forecasting. Our key insight is that *the missing signal lies not only in the predicted future states themselves, but more importantly in the transitions between them*. As illustrated in Figure 1(c), inverse dynamics naturally decodes this signal. While forward world modeling predicts how the scene evolves under a motion hypothesis, inverse dynamics asks the complementary question of *what motion-relevant change is implied by the transition from one state to the next*. In latent Bird's Eye View space, adjacent predicted futures capture the expected evolution of the ego vehicle and surrounding scene under a planning hypothesis, and their transitions therefore contain recoverable motion semantics reflecting trajectory adjustments consistent with that evolution. Rather than treating future latent states as static planning features, we decode adjacent latent transitions with an inverse dynamics model to extract transition-aware trajectory features and infer motion deltas for planning optimization. In this way, inverse dynamics explicitly bridges world modeling and trajectory generation, transforming future prediction from *passive scene forecasting* into *planning-aware transition reasoning*.

Based on this perspective, we propose **IDOL**, a world-model-based end-to-end planner operating in latent BEV space that tightly couples future reasoning with trajectory optimization. IDOL first predicts multiple future latent scene states with a BEV world model, providing anticipated scene evolution beyond the current observation. It then applies an inverse dynamics model to adjacent latent futures to infer motion-aware transition features and planning-relevant trajectory deltas that explain the predicted evolution. These signals are aggregated and injected into the planner to optimize the anchor trajectory, allowing motion updates to follow the dynamics implied by future state transitions rather than future context alone. A lightweight closed-loop refinement module further re-evaluates the updated trajectory to improve long-horizon consistency.

In summary, our contributions are as follows:

i. We identify a fundamental gap in recent end-to-end and world-model-based planners: current methods still lack a sufficiently explicit and interpretable mechanism for converting predicted future evolution into planning-relevant trajectory updates, which limits the actionability of future reasoning for executable motion generation.

ii. We propose IDOL, a world-model-based end-to-end planner in latent BEV space that introduces inverse dynamics into latent future reasoning, thereby establishing an explicit bridge between world modeling and executable motion generation. We further design a lightweight closed-loop refinement module to improve long-horizon consistency through another round of future-aware reasoning.

iii. We conduct extensive experiments on both NAVSIM v1 and NAVSIM v2, including comprehensive comparisons and analysis, and show that our method consistently achieves state-of-the-art performance among comparable methods. These results validate the effectiveness of our design in improving planning quality across different benchmark settings.

2 Related work

2.1 End-to-end autonomous driving

End-to-end autonomous driving (Zhang et al., 2025a; Li et al., 2025g; Feng et al., 2025a; Li & Cui, 2024; Luo et al., 2025; Chen et al., 2024a; Feng et al., 2025b; Shang et al., 2026) directly maps sensor observations to waypoints, trajectories, or low-level controls, reducing error accumulation and handcrafted interfaces in modular pipelines. Early methods (Chitta et al., 2022) largely relied on imitation learning, while planning-oriented formulations (Guo et al., 2025; Wozniak et al., 2025; Yao et al., 2026; Hwang et al., 2024; Huang et al., 2025; Sun et al., 2025a; Li et al., 2025a; Sun et al., 2026a) later became dominant. TransFuser (Chitta et al., 2022) established a strong transformer-based sensor-fusion baseline, while UniAD (Hu et al., 2023b) unified perception and prediction around downstream planning. DriveMamba (Su et al., 2026) improved efficiency and scalability with a task-centric state-space architecture, and LEAD (Nguyen et al., 2025) revisited imitation learning through learner-expert asymmetry. DiffusionDrive (Liao et al., 2025) and GoalFlow (Xing et al., 2025) introduced generative planning via truncated diffusion and flow matching to better model multi-modal trajectories. More recently, VLA-based end-to-end driving (Zheng et al., 2025a; Rawal et al., 2026; Xu et al., 2025; Wang et al., 2026a) has extended this direction toward language-grounded action generation and post-training alignment. DriveVLA-W0 (Li et al., 2025d) incorporated world-model objectives to provide dense supervision and improve data scaling, while AutoDrive-R^2 (Yuan et al., 2026) introduced chain-of-thought reasoning, self-reflection, and GRPO-based post-training for physically plausible trajectory generation. Despite these advances, most end-to-end planners still couple planning mainly to the current scene or output-level multimodality, leaving the connection between future state transitions and action generation largely implicit.

2.2 World model in autonomous driving

World models in autonomous driving (Chen et al., 2025; Wang et al., 2024a; Li et al., 2024a; Jiang et al., 2025c) predict scene dynamics for forecasting, planning, simulation, and long-horizon decisions. Dreamer (Hafner et al., 2019; 2020; 2023) introduced latent imagined rollouts for long-horizon planning. GAIA-1 (Hu et al., 2023a) formulated driving world modeling as action-conditioned multimodal sequence generation, while Vista (Gao et al., 2024) emphasized environmental diversity and controllability. WoTE (Li et al., 2025e) evaluates trajectory candidates through predicted BEV states. SeerDrive (Zhang et al., 2025a) couples scene evolution and planning through iterative bidirectional refinement. DriveLaW (Xia et al., 2025) unifies video generation and motion planning in shared latents that support closed-loop planning.

Despite this progress, the key challenge is how predicted futures shape decisions. Although recent methods (Wang et al., 2024b; Zheng et al., 2025b; Zhang et al., 2025a) jointly train world modeling and planning, the connection between future scene dynamics and planned motion remains insufficiently explained. We address this gap by making latent BEV transitions action-relevant, constraining planning through both anticipated scene evolution and its implied control signals.

2.3 Inverse dynamics models

Inverse dynamics models (Hollerbach, 1980; Otten, 2003; Koopman et al., 1995; Angeles et al., 1989) originate from physics, control, and reinforcement learning, where the goal is to infer the action or control input that caused a transition between two states. This viewpoint is useful because it makes action semantics explicit: an inverse dynamics model can identify which aspects of a

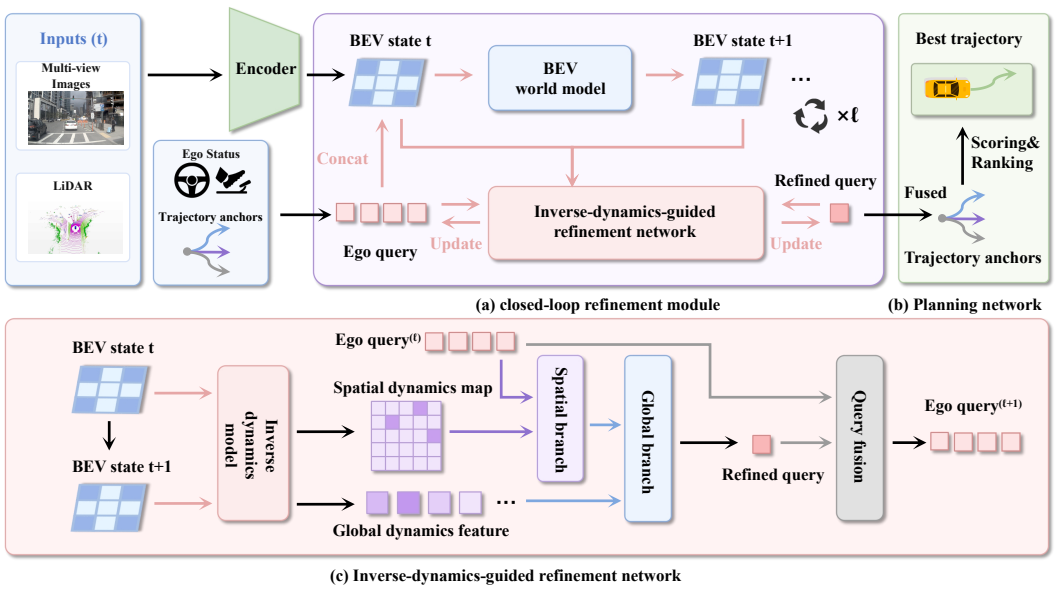


Figure 2: Overview of the proposed IDOL framework. (a) The closed-loop refinement module rolls out future latent BEV states with the BEV world model and refines the planning query through inverse-dynamics feedback. (b) The planning network fuses the refined query with trajectory anchors, predicts candidate trajectories, and selects the best one by scoring and ranking. (c) The inverse-dynamics-guided refinement network decodes adjacent future BEV states into spatial and global dynamics cues, which are fused to update the ego query.

transition are action-relevant, recover latent controls when actions are unavailable or noisy, and encourage representations that preserve controllable structure rather than merely appearance. Recent machine learning work (Tian et al., 2024) has employed inverse dynamics for model-based control, representation learning, and general sequential decision-making.

In autonomous driving, inverse dynamics remains comparatively underexplored and is mostly used as an auxiliary component. ReSim (Yang et al., 2025) employs an inverse dynamics model to convert predicted videos into executable trajectories and assess controllability, while FutureSightDrive (Zeng et al., 2025) formulates planning as recovering actions from current observations and imagined future scenes. These works suggest that inverse dynamics can bridge predicted future states and executable actions. However, existing driving world models typically focus on latent future forecasting, trajectory evaluation, or reward-guided selection, without explicitly learning transition-to-action mappings between adjacent predicted states. Our method addresses this gap by introducing an inverse dynamics module over predicted latent BEV states, explicitly coupling state evolution with planning and aligning trajectory generation with action-consistent future transitions.

3 Methodology

3.1 Preliminary

End-to-end autonomous driving. End-to-end autonomous driving formulates trajectory planning as learning a direct mapping from raw sensory observations and ego-state inputs to future ego motion, without relying on a manually designed modular stack. Recent world-model-based methods further improve this paradigm by explicitly modeling future scene evolution and using the predicted futures to guide planning.

Inverse dynamics model. Inverse dynamics model infers the latent action or motion representation that explains the transition between two adjacent states. In its general form, given two consec-

utive states ξ_t and ξ_{t+1}, inverse dynamics seeks a compact transition descriptor

$$d_t = f_{\text{idm}}(\xi_t, \xi_{t+1}), \tag{1}$$

which captures the motion-consistent change from ξ_t to ξ_{t+1}. This formulation is particularly suitable for autonomous driving, since adjacent future states naturally encode how ego motion interacts with scene evolution. In our setting, the two states are imagined BEV latent states produced by the world model. Therefore, inverse dynamics provides a natural bridge from *future state transition* to *trajectory-related motion features*, making it well suited for refining trajectories from consecutive imagined BEV frames.

3.2 Scene representation

As shown in Figure 2, we encode the current scene into a latent BEV representation B using a TransFuser-style image-LiDAR backbone (Chitta et al., 2022). We adopt an offline trajectory-anchor vocabulary $\mathcal{A} = \{\tau_k\}_{k=1}^{K}$, where $\tau_k \in \mathbb{R}^{S\times 3}$ represents a candidate future motion (Chen et al., 2024b; Li et al., 2024b). Each anchor feature a_k is concatenated with the ego-state feature z to initialize a candidate-specific planning query:

$$e_{k,\text{init}} = \phi([z, a_k]), \qquad k = 1, \ldots, K, \tag{2}$$

which guides subsequent future imagination and trajectory refinement.

3.3 Latent BEV world model

For each trajectory candidate k and refinement iteration ℓ, we initialize the latent rollout with $q_{k,\ell}^{(0)} = e_{k,\text{in}}^{(\ell)}$ and $B_{k,\ell}^{(0)} = B$. At each step u, the ego query is injected at the corresponding future location of anchor τ_k, and a transformer-based world model predicts the next query and latent BEV state:

$$B_{k,\ell,\text{cond}}^{(u)} = \mathcal{I}\Big(B_{k,\ell}^{(u)}, q_{k,\ell}^{(u)}; \tau_k\Big), \tag{3}$$

$$\Big[q_{k,\ell}^{(u+1)}, B_{k,\ell}^{(u+1)}\Big] = \text{BEVWorldModel}\Big(\Big[q_{k,\ell}^{(u)}, B_{k,\ell,\text{cond}}^{(u)}\Big]\Big), \tag{4}$$

where $\mathcal{I}$ denotes query injection. Repeating this process with the same anchor yields the candidate-specific rollout

$$\mathcal{R}_k^{(\ell)} = \Big\{\Big(q_{k,\ell}^{(u)}, B_{k,\ell,\text{cond}}^{(u)}\Big)\Big\}_{u=0}^{U}, \tag{5}$$

whose adjacent candidate-conditioned BEV states are subsequently decoded by the inverse dynamics model.

3.4 Inverse-dynamics-guided refinement network

3.4.1 Inverse dynamics model

While the world model predicts future latent scene evolution, planning still requires a mechanism that translates the imagined transitions into motion-relevant feedback. To bridge this gap, we introduce an inverse dynamics model (IDM).

For each pair of adjacent candidate-conditioned latent BEV states within the ℓ-th rollout, the IDM produces two complementary transition representations:

$$\Big(S_{k,\ell}^{(u)}, g_{k,\ell}^{(u)}\Big) = \text{IDM}\Big(B_{k,\ell,\text{cond}}^{(u)}, B_{k,\ell,\text{cond}}^{(u+1)}\Big), \tag{6}$$

where $S_{k,\ell}^{(u)} \in \mathbb{R}^{H\times W\times C}$ is a spatial dynamics map that preserves position-wise motion variations across the latent BEV feature map, and $g_{k,\ell}^{(u)} \in \mathbb{R}^{C}$ is a pooled global dynamics feature that summarizes the overall transition.

When multiple rollout transitions are available, we aggregate the transition representations over the rollout horizon:

$$S_k^{(\ell)} = \frac{1}{U}\sum_{u=0}^{U-1} S_{k,\ell}^{(u)}, \qquad g_k^{(\ell)} = \frac{1}{U}\sum_{u=0}^{U-1} g_{k,\ell}^{(u)}. \tag{7}$$

The key idea is that future transitions should capture not only *what* changes globally, but also *where* critical changes occur. A pooled inverse-dynamics feature alone is often too coarse, since only part of the spatial variation is relevant to the current trajectory hypothesis. The spatial representation preserves localized transition evidence, while the global representation summarizes overall evolution. Together, they bridge latent future prediction and motion refinement at the transition level.

3.4.2 CLOSED-LOOP QUERY REFINEMENT

We refine the planning query by combining the spatial and global inverse-dynamics cues through a dual-branch mechanism.

For the first closed-loop iteration, the refinement process is initialized by the anchor-conditioned query:

$$e_{k,\mathrm{in}}^{(0)} = e_{k,\mathrm{init}}. \tag{8}$$

Given the transition representations extracted from the corresponding latent rollout, the spatial branch first allows the current planning query to retrieve candidate-relevant local motion evidence through cross-attention:

$$e_{k,\mathrm{spa}}^{(\ell)} = \mathrm{LN}\Big(e_{k,\mathrm{in}}^{(\ell)} + \mathrm{CrossAttn}\Big(e_{k,\mathrm{in}}^{(\ell)}, S_k^{(\ell)}, S_k^{(\ell)}\Big)\Big). \tag{9}$$

Here, $e_{k,\mathrm{spa}}^{(\ell)}$ denotes the query after spatial transition-aware refinement.

The global dynamics representation is subsequently used to calibrate the spatially refined query through an MLN (Wang et al., 2023):

$$e_{k,\mathrm{ref}}^{(\ell)} = \mathrm{MLN}_{\mathrm{idm}}\Big(e_{k,\mathrm{spa}}^{(\ell)}, g_k^{(\ell)}\Big). \tag{10}$$

This dual-branch design assigns complementary roles to the two inverse-dynamics representations: the spatial branch determines *where* motion-relevant future changes occur, while the global branch provides holistic transition-aware calibration. Consequently, the feedback from imagined futures becomes more spatially selective and informative than using a pooled inverse-dynamics representation alone.

To prevent iterative drift, subsequent closed-loop iterations re-anchor the refined query to the original anchor-conditioned query:

$$e_{k,\mathrm{in}}^{(\ell+1)} = \mathrm{MLN}_{\mathrm{cl}}\Big(e_{k,\mathrm{init}}, e_{k,\mathrm{ref}}^{(\ell)}\Big), \qquad \ell = 0, \ldots, L-2. \tag{11}$$

The updated query $e_{k,\mathrm{in}}^{(\ell+1)}$ initializes the next candidate-conditioned rollout, while the anchor τ_k remains fixed. Trajectory offsets are decoded only after the final refinement, where $e_{k,\mathrm{ref}}^{(L-1)}$ is fed into a transformer decoder and an MLP head. The resulting candidates are ranked by a reward model (Li et al., 2025e) to select the final trajectory.

3.5 TRAINING LOSS

Our network is trained with trajectory offset regression, imitation- and simulator-reward supervision, and current/future BEV semantic supervision. We additionally supervise motion correction in the final global IDM feature:

$$\mathcal{L}_{\mathrm{corr}} = \ell_{\mathrm{SL1}}\Big(h_{\mathrm{mot}}(g_{k^\star}^{(L-1)}), \mathrm{sg}\big[E_{\mathrm{mot}}(\tau^{\mathrm{GT}}) - E_{\mathrm{mot}}(\tau_{k^\star})\big]\Big), \tag{12}$$

where $k^\star$ is the closest anchor under winner-take-all assignment, E_{mot} concatenates fixed normalized speed, acceleration, and heading-rate sequences, and sg denotes stop-gradient. The masked Smooth L1 loss ℓ_{SL1} averages errors within each motion component and then across components, excluding unreliable low-speed heading rates. The head h_{mot} is used only during training.

The overall objective is

$$\mathcal{L} = \lambda_{\mathrm{off}}\mathcal{L}_{\mathrm{off}} + \lambda_{\text{off-im}}\mathcal{L}_{\text{off-im}} + \lambda_{\mathrm{im}}\mathcal{L}_{\mathrm{im}} + \lambda_{\mathrm{sim}}\mathcal{L}_{\mathrm{sim}} + \lambda_{\mathrm{map}}\mathcal{L}_{\mathrm{map}} + \lambda_{\mathrm{corr}}\mathcal{L}_{\mathrm{corr}}, \tag{13}$$

where the first five terms supervise offsets, auxiliary and main imitation rewards, simulator metrics, and weighted current/future BEV semantics, respectively. Loss definitions and coefficients are detailed in Appendix B.

Table 1: Performance comparison on the NAVSIM v1 navtest split with closed-loop metrics. "C" and "L" denote camera and LiDAR inputs, respectively. For fair comparison, all listed results are reported under the ResNet-34 image-backbone setting. Bold and underlined values indicate the best and second-best results, respectively. ResWorld* denotes the historical-frame variant that uses historical frames to obtain temporal residuals, and DiffE2E† denotes the variant with a ResNet-34 image encoder.

Method	Img. Backbone	Input	NC ↑	DAC ↑	TTC ↑	Comf. ↑	EP ↑	PDMS ↑
VADv2 (Chen et al., 2024b)	ResNet-34	C & L	97.2	89.1	91.6	**100**	76.0	80.9
UniAD (Hu et al., 2023b)	ResNet-34	Camera	97.8	91.9	92.9	**100**	78.8	83.4
PARA-Drive (Weng et al., 2024)	ResNet-34	Camera	97.9	92.4	93.0	99.8	79.3	84.0
TransFuser (Chitta et al., 2022)	ResNet-34	C & L	97.7	92.8	92.8	**100**	79.2	84.0
LAW (Li et al., 2024a)	ResNet-34	Camera	96.4	95.4	88.7	<u>99.9</u>	81.7	84.6
DiffusionDrive (Liao et al., 2025)	ResNet-34	C & L	98.2	96.2	94.7	**100**	82.2	88.1
WoTE (Li et al., 2025e)	ResNet-34	C & L	98.5	96.8	94.9	<u>99.9</u>	81.9	88.3
SeerDrive (Zhang et al., 2025a)	ResNet-34	C & L	98.4	<u>97.0</u>	94.9	<u>99.9</u>	83.2	88.9
ResWorld* (Zhang et al., 2026)	ResNet-34	C & L	<u>98.9</u>	96.5	<u>95.6</u>	**100**	83.1	89.0
MeanFuser (Wang et al., 2026b)	ResNet-34	Camera	98.6	<u>97.0</u>	95.0	**100**	82.8	89.0
DiffE2E† (Zhao et al., 2025)	ResNet-34	C & L	**99.2**	96.8	**96.7**	**100**	<u>83.6</u>	<u>89.8</u>
IDOL	ResNet-34	C & L	98.7	**97.7**	95.5	**100**	**84.0**	**90.0**

4 EXPERIMENTS

4.1 BENCHMARK

We evaluate our method on NAVSIM v1 (Dauner et al., 2024) and NAVSIM v2 (Cao et al., 2025) under the official closed-loop protocols. On NAVSIM v1, we report results on the navtest split using the Predictive Driver Model Score (PDMS), which aggregates no-collision (NC), drivable area compliance (DAC), time-to-collision (TTC), comfort, and ego progress (EP). On NAVSIM v2, we further evaluate on the challenging navhard split under the two-stage pseudo-simulation protocol using the Extended PDMS (EPDMS), which additionally considers driving direction compliance (DDC), traffic light compliance (TLC), lane keeping (LK), history comfort (HC), and extended comfort (EC).

4.2 IMPLEMENTATION DETAILS

We use a planning horizon of 4 s with a sampling interval of 0.5 s, resulting in 8 future waypoints. Our model adopts a ResNet34-based TransFuser backbone for multi-modal fusion, where the stitched front-view image input is resized to 256×1024. The latent feature dimension is set to 256. In the current implementation, the BEV representation is flattened into 64 latent queries, with $H_{\text{BEV}} \times W_{\text{BEV}} = 8 \times 8$, and trajectory planning is performed over 256 trajectory anchors. The resulting model contains 69.36M parameters and achieves 17.65 FPS on a single NVIDIA RTX 3090 GPU with batch size 1, indicating that IDOL maintains practical inference efficiency while introducing inverse-dynamics-guided future reasoning.

Training is conducted on 4 NVIDIA GeForce RTX 3090 GPUs with a batch size of 4 per GPU for approximately 24 hours. We use AdamW with an initial learning rate of 2×10^{-4} and weight decay of 1×10^{-4}.

4.3 COMPARISON WITH SOTA

Table 1 compares IDOL with recent end-to-end planners on the NAVSIM navtest split using a ResNet-34 (He et al., 2016) image backbone. IDOL achieves the best PDMS of 90.0. Table 2 further evaluates the challenging NAVSIM v2 navhard split, where IDOL attains the highest final EPDMS of 38.0, outperforming the strongest baseline WoTE by 10.1 points.

Beyond the main comparisons, Table 9 reports the extended NAVSIM v2 navtest benchmark. IDOL achieves 89.6 EPDMS, ranking first among learned methods and remaining competitive with the human agent. Table 10 further evaluates the more challenging navhard stage-1-only setting, where

Table 2: Performance comparison on the NAVSIM v2 navhard split. PDM-Closed is listed separately as a privileged planner using ground-truth perception.

Method	Backbone	Stage	NC ↑	DAC ↑	DDC ↑	TLC ↑	EP ↑	TTC ↑	LK ↑	HC ↑	EC ↑	EPDMS ↑
PDM-Closed (Dauner et al., 2023)	-	Stage 1	94.4	98.8	100.0	99.5	100.0	93.5	99.3	87.7	36.0	51.3
		Stage 2	88.1	90.6	96.3	98.5	100.0	83.1	73.7	91.5	25.4	
LTF (Prakash et al., 2021)	ResNet34	Stage 1	96.2	79.5	**99.1**	99.5	**84.1**	95.1	94.2	97.5	**79.1**	23.1
		Stage 2	77.7	70.2	84.2	98.0	85.1	75.6	45.4	95.7	**75.9**	
GTRS-DP (Li et al., 2025h)	ResNet34	Stage 1	94.7	78.8	96.1	99.5	83.0	94.4	92.0	97.5	72.8	23.8
		Stage 2	80.3	74.4	84.9	98.0	81.9	78.8	45.4	**96.7**	70.1	
DiffusionDrive (Liao et al., 2025)	ResNet34	Stage 1	96.0	79.7	97.4	99.5	81.3	93.1	90.8	96.8	73.8	24.2
		Stage 2	82.1	72.2	88.5	98.7	85.1	78.8	49.2	89.3	71.2	
GuideFlow (Liu et al., 2025a)	ResNet34	Stage 1	96.6	80.5	96.3	99.3	82.3	94.9	91.5	**97.7**	67.8	27.1
		Stage 2	**87.3**	76.7	**88.8**	**99.2**	84.3	**85.1**	49.7	93.1	44.5	
WoTE (Li et al., 2025e)	ResNet34	Stage 1	**97.4**	88.2	97.7	99.3	82.7	96.4	90.8	97.3	68.0	27.9
		Stage 2	81.2	77.7	84.8	98.1	**85.9**	78.5	46.2	96.6	63.3	
IDOL	ResNet34	Stage 1	97.2	**89.6**	98.0	**99.6**	82.3	**96.9**	**95.3**	97.6	72.9	**38.0**
		Stage 2	84.9	**81.7**	87.7	98.5	84.7	81.1	**50.0**	95.4	64.0	

IDOL achieves the best performance with 76.2 EPDMS, surpassing Hydra-MDP† by 3.1 points and WoTE by 9.5 points. Across NAVSIM v1 navtest, NAVSIM v2 navtest, navhard stage-1, and the two-stage navhard evaluation, IDOL consistently achieves state-of-the-art performance among comparable learned planners, demonstrating the robustness and stability of inverse-dynamics-guided future reasoning under both standard and challenging long-tail driving scenarios.

4.4 Ablation Study

We conduct ablation studies on the NAVSIM navtest split to analyze the contribution of each design in IDOL. We report PDMS and its five component metrics. Additional ablation results and analyses are provided in Appendix E.

Effect of inverse dynamics and closed-loop refinement. Table 3 studies the main components of IDOL. Without IDM, future BEV states lack explicit transition-to-motion feedback, limiting their effect on trajectory refinement. Introducing IDM improves PDMS from 87.3 to 89.2 by converting latent future transitions into motion-aware query updates. A single closed-loop refinement iteration further raises PDMS to 90.0 by feeding the updated query back into future reasoning. Increasing the number of iterations to two or three reduces performance, suggesting possible over-correction.

Table 3: Main ablation of inverse dynamics and closed-loop refinement on the NAVSIM navtest split. IDM denotes the inverse dynamics model, and CL denotes closed-loop refinement.

IDM	CL iters.	NC ↑	DAC ↑	TTC ↑	Comf. ↑	EP ↑	PDMS ↑
×	×	98.5	95.0	95.2	**100**	81.5	87.3
✓	×	98.6	97.0	95.4	**100**	83.3	89.2
✓	1	98.7	**97.7**	95.5	**100**	**84.0**	**90.0**
✓	2	98.6	97.3	95.2	**100**	83.5	89.4
✓	3	**98.8**	97.2	**95.7**	**100**	83.1	89.4

Temporal design of IDM. Table 4 compares different temporal inputs for IDM. The two-frame design directly decodes adjacent BEV transitions, while the four-frame variant uses a longer window that may dilute local transition cues for immediate refinement. The better performance of the two-frame setting shows that adjacent transitions provide a more compact and effective basis for planning-oriented motion refinement.

Effect of spatial and global dynamics branches. Table 5 analyzes the dual-branch design. The global branch captures holistic transition patterns but lacks spatial selectivity, while the spatial branch preserves local transition evidence but lacks global calibration. Combining both branches

Table 4: Ablation of temporal decoding in the inverse dynamics model. The two-frame variant decodes adjacent BEV transitions, while the four-frame variant uses a longer future window.

IDM temporal input	NC ↑	DAC ↑	TTC ↑	Comf. ↑	EP ↑	PDMS ↑
4 frame	**98.7**	97.3	**95.6**	**100**	83.7	89.6
2 frame	**98.7**	**97.7**	95.5	**100**	**84.0**	**90.0**

provides spatially selective and globally calibrated future-to-planning feedback, achieving the best overall PDMS.

Table 5: Ablation of the spatial and global dynamics branches in the inverse-dynamics-guided refinement network.

Dynamics Branch	NC ↑	DAC ↑	TTC ↑	Comf. ↑	EP ↑	PDMS ↑
w/o spatial branch	**98.8**	96.9	**95.8**	**100**	83.2	89.3
w/o global branch	98.7	97.2	95.3	**100**	83.9	89.5
dual-branch IDM	98.7	**97.7**	95.5	**100**	**84.0**	**90.0**

4.5 Visualization

To better understand inverse-dynamics-guided refinement, we visualize a representative NAVSIM navtest case in Figure 3. Panels (a)–(c) show the latent BEV change, the corresponding IDM spatial response, and the refinement-query attention, respectively. The white contour marks high-change regions extracted from Panel (a). The visualization shows that IDM indeed concentrates on the motion-informative parts within the latent BEV change regions, rather than unrelated background variations, and the query attends to these motion-aware cues during refinement. Panel (d) compares the original, refined, and expert trajectories, where the refined trajectory better aligns with the expert, suggesting that IDM converts latent future transitions into actionable planning updates. More qualitative results are provided in Appendix F.

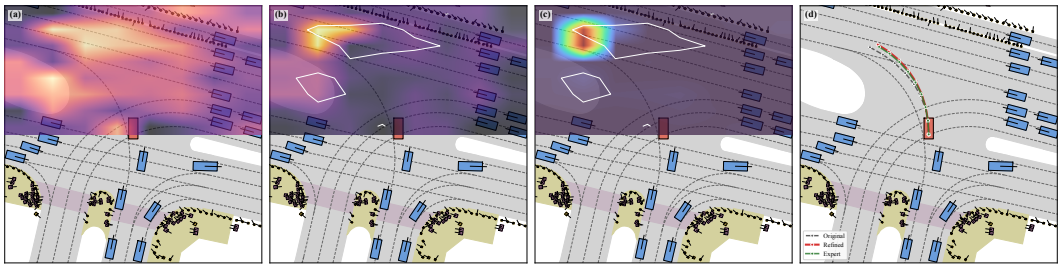


Figure 3: Qualitative visualization of inverse-dynamics-guided refinement on the NAVSIM navtest split.

5 Conclusion

In this work, we proposed IDOL, an inverse-dynamics-guided future prediction framework for end-to-end autonomous driving. By decoding adjacent latent BEV futures, IDOL converts predicted scene evolution into transition-aware motion cues and refines trajectories through a lightweight closed-loop process. This explicitly bridges world modeling and executable motion generation, making future prediction more actionable for planning. Experiments on NAVSIM v1 and NAVSIM v2 show that IDOL achieves state-of-the-art performance on navtest, navhard stage-1, and two-stage navhard benchmarks, demonstrating the effectiveness and generality of inverse-dynamics-guided future reasoning. This framework provides a general paradigm for integrating future state transitions with motion generation in future autonomous driving research.

AI USE STATEMENT

We used generative AI tools to assist with drafting and editing parts of the manuscript, improving language and readability, suggesting paper structure, searching for and identifying relevant literature, and editing a portion of software code. All AI-assisted text was reviewed and revised by the authors, all cited literature and factual claims were manually verified against the original sources, and AI-assisted code edits were reviewed and tested for correctness.

Generative AI tools were not used to originate the core research idea, formulate or refine the research hypotheses, design or provide feedback on the proposed methodology or experiments, implement the proposed method, generate or process experimental data, derive theoretical or mathematical claims, or interpret experimental results. Other tasks requiring disclosure under the ICLR 2027 AI Policy were not applicable to this work.

The authors take full responsibility for the final content of this work, including all text, claims, code, and other artifacts produced with the assistance of generative AI.

REPRODUCIBILITY STATEMENT

We provide the necessary information to facilitate reproduction of our results. The architecture and formulation of IDOL, including the latent BEV world model, inverse dynamics module, trajectory refinement, and training objectives, are described in the main paper. The experimental section specifies the benchmark splits, evaluation metrics, model configuration, training settings, and inference setup used in our experiments. Additional implementation analyses and controlled ablation studies are provided in the appendix. All results on NAVSIM v1 and NAVSIM v2 follow the official evaluation protocols and benchmark settings. Anonymous source code and configuration files are provided in the supplementary material to further facilitate reproducibility.

REFERENCES

J Angeles, Ou Ma, and A Rojas. An algorithm for the inverse dynamics of n-axis general manipulators using kane's equations. *Computers & Mathematics with Applications*, 17(12):1545–1561, 1989.

Holger Caesar, Varun Bankiti, Alex H Lang, Sourabh Vora, Venice Erin Liong, Qiang Xu, Anush Krishnan, Yu Pan, Giancarlo Baldan, and Oscar Beijbom. nuscenes: A multimodal dataset for autonomous driving. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 11621–11631, 2020.

Wei Cao, Marcel Hallgarten, Tianyu Li, Daniel Dauner, Xunjiang Gu, Caojun Wang, Yakov Miron, Marco Aiello, Hongyang Li, Igor Gilitschenski, et al. Pseudo-simulation for autonomous driving. *arXiv preprint arXiv:2506.04218*, 2025.

Li Chen, Penghao Wu, Kashyap Chitta, Bernhard Jaeger, Andreas Geiger, and Hongyang Li. End-to-end autonomous driving: Challenges and frontiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(12):10164–10183, 2024a.

Shaoyu Chen, Bo Jiang, Hao Gao, Bencheng Liao, Qing Xu, Qian Zhang, Chang Huang, Wenyu Liu, and Xinggang Wang. Vadv2: End-to-end vectorized autonomous driving via probabilistic planning. *arXiv preprint arXiv:2402.13243*, 2024b.

Yuntao Chen, Yuqi Wang, and Zhaoxiang Zhang. Drivinggpt: Unifying driving world modeling and planning with multi-modal autoregressive transformers. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 26890–26900, 2025.

Kashyap Chitta, Aditya Prakash, Bernhard Jaeger, Zehao Yu, Katrin Renz, and Andreas Geiger. Transfuser: Imitation with transformer-based sensor fusion for autonomous driving. *IEEE transactions on pattern analysis and machine intelligence*, 45(11):12878–12895, 2022.

Daniel Dauner, Marcel Hallgarten, Andreas Geiger, and Kashyap Chitta. Parting with misconceptions about learning-based vehicle motion planning. In *Conference on Robot Learning*, pp. 1268–1281. PMLR, 2023.

Daniel Dauner, Marcel Hallgarten, Tianyu Li, Xinshuo Weng, Zhiyu Huang, Zetong Yang, Hongyang Li, Igor Gilitschenski, Boris Ivanovic, Marco Pavone, et al. Navsim: Data-driven non-reactive autonomous vehicle simulation and benchmarking. *Advances in Neural Information Processing Systems*, 37:28706–28719, 2024.

Lan Feng, Yang Gao, Eloi Zablocki, Quanyi Li, Wuyang Li, Sichao Liu, Matthieu Cord, and Alexandre Alahi. Rap: 3d rasterization augmented end-to-end planning. *arXiv preprint arXiv:2510.04333*, 2025a.

Renju Feng, Ning Xi, Duanfeng Chu, Rukang Wang, Zejian Deng, Anzheng Wang, Liping Lu, Jinxiang Wang, and Yanjun Huang. Artemis: Autoregressive end-to-end trajectory planning with mixture of experts for autonomous driving. *IEEE Robotics and Automation Letters*, 11(1):226–233, 2025b.

Shenyuan Gao, Jiazhi Yang, Li Chen, Kashyap Chitta, Yihang Qiu, Andreas Geiger, Jun Zhang, and Hongyang Li. Vista: A generalizable driving world model with high fidelity and versatile controllability. *Advances in Neural Information Processing Systems*, 37:91560–91596, 2024.

Ke Guo, Haochen Liu, Xiaojun Wu, Jia Pan, and Chen Lv. ipad: Iterative proposal-centric end-to-end autonomous driving. *arXiv preprint arXiv:2505.15111*, 2025.

Mingzhe Guo, Yixiang Yang, Chuanrong Han, Rufeng Zhang, Shirui Li, Ji Wan, and Zhipeng Zhang. Flowad: Ego-scene interactive modeling for autonomous driving. *arXiv preprint arXiv:2603.13399*, 2026.

Danijar Hafner, Timothy Lillicrap, Jimmy Ba, and Mohammad Norouzi. Dream to control: Learning behaviors by latent imagination. *arXiv preprint arXiv:1912.01603*, 2019.

Danijar Hafner, Timothy Lillicrap, Mohammad Norouzi, and Jimmy Ba. Mastering atari with discrete world models. *arXiv preprint arXiv:2010.02193*, 2020.

Danijar Hafner, Jurgis Pasukonis, Jimmy Ba, and Timothy Lillicrap. Mastering diverse domains through world models. *arXiv preprint arXiv:2301.04104*, 2023.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

John M Hollerbach. A recursive lagrangian formulation of maniputator dynamics and a comparative study of dynamics formulation complexity. *IEEE Transactions on Systems, Man, and Cybernetics*, 10(11):730–736, 1980.

Anthony Hu, Lloyd Russell, Hudson Yeo, Zak Murez, George Fedoseev, Alex Kendall, Jamie Shotton, and Gianluca Corrado. Gaia-1: A generative world model for autonomous driving. *arXiv preprint arXiv:2309.17080*, 2023a.

Yihan Hu, Jiazhi Yang, Li Chen, Keyu Li, Chonghao Sima, Xizhou Zhu, Siqi Chai, Senyao Du, Tianwei Lin, Wenhai Wang, et al. Planning-oriented autonomous driving. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 17853–17862, 2023b.

Zhao Huang, Bowen Zhang, Zhongzhu Li, and Di Lin. Occdriver: Future occupancy guided dual-branch trajectory planner in autonomous driving. In *International Conference on Learning Representations*, volume 2026, pp. 141940–141962, 2026.

Zhiyu Huang, Xinshuo Weng, Maximilian Igl, Yuxiao Chen, Yulong Cao, Boris Ivanovic, Marco Pavone, and Chen Lv. Gen-drive: Enhancing diffusion generative driving policies with reward modeling and reinforcement learning fine-tuning. In *2025 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3445–3451. IEEE, 2025.

Jyh-Jing Hwang, Runsheng Xu, Hubert Lin, Wei-Chih Hung, Jingwei Ji, Kristy Choi, Di Huang, Tong He, Paul Covington, Benjamin Sapp, et al. Emma: End-to-end multimodal model for autonomous driving. *arXiv preprint arXiv:2410.23262*, 2024.

Anqing Jiang, Yu Gao, Zhigang Sun, Yiru Wang, Jijun Wang, Jinghao Chai, Qian Cao, Yuweng Heng, Hao Jiang, Yunda Dong, et al. Diffvla: Vision-language guided diffusion planning for autonomous driving. *arXiv preprint arXiv:2505.19381*, 2025a.

Anqing Jiang, Yu Gao, Yiru Wang, Zhigang Sun, Shuo Wang, Yuwen Heng, Hao Sun, Shichen Tang, Lijuan Zhu, Jinhao Chai, et al. Irl-vla: Training an vision-language-action policy via reward world model. *arXiv preprint arXiv:2508.06571*, 2025b.

Bo Jiang, Shaoyu Chen, Qing Xu, Bencheng Liao, Jiajie Chen, Helong Zhou, Qian Zhang, Wenyu Liu, Chang Huang, and Xinggang Wang. Vad: Vectorized scene representation for efficient autonomous driving. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 8340–8350, 2023.

Guangfeng Jiang, Yueru Luo, Jun Liu, Yi Huang, Yiyao Zhu, Zhan Qu, Dave Zhenyu Chen, Bingbing Liu, and Xu Yan. Wpt: World-to-policy transfer via online world model distillation. *arXiv preprint arXiv:2511.20095*, 2025c.

Jungho Kim, Jiyong Oh, Seunghoon Yu, Hongjae Shin, Donghyuk Kwak, and Jun Won Choi. Safedrive: Fine-grained safety reasoning for end-to-end driving in a sparse world. *arXiv preprint arXiv:2602.18887*, 2026.

Bart Koopman, Henk J Grootenboer, and Henk J De Jongh. An inverse dynamics model for the analysis, reconstruction and prediction of bipedal walking. *Journal of biomechanics*, 28(11): 1369–1376, 1995.

Jingyu Li, Bozhou Zhang, Xin Jin, Jiankang Deng, Xiatian Zhu, and Li Zhang. Imagidrive: A unified imagination-and-planning framework for autonomous driving. *arXiv preprint arXiv:2508.11428*, 2025a.

Jingyu Li, Junjie Wu, Dongnan Hu, Xiangkai Huang, Bin Sun, Zhihui Hao, Xianpeng Lang, Xiatian Zhu, and Li Zhang. Sgdrive: Scene-to-goal hierarchical world cognition for autonomous driving. *arXiv preprint arXiv:2601.05640*, 2026.

Kailin Li, Zhenxin Li, Shiyi Lan, Yuan Xie, Zhizhong Zhang, Jiayi Liu, Zuxuan Wu, Zhiding Yu, and Jose M Alvarez. Hydra-mdp++: Advancing end-to-end driving via expert-guided hydra-distillation. *arXiv preprint arXiv:2503.12820*, 2025b.

Peidong Li and Dixiao Cui. Navigation-guided sparse scene representation for end-to-end autonomous driving. *arXiv preprint arXiv:2409.18341*, 2024.

Pengxiang Li, Yinan Zheng, Yue Wang, Huimin Wang, Hang Zhao, Jingjing Liu, Xianyuan Zhan, Kun Zhan, and Xianpeng Lang. Discrete diffusion for reflective vision-language-action models in autonomous driving. *arXiv preprint arXiv:2509.20109*, 2025c.

Yingyan Li, Lue Fan, Jiawei He, Yuqi Wang, Yuntao Chen, Zhaoxiang Zhang, and Tieniu Tan. Enhancing end-to-end autonomous driving with latent world model. *arXiv preprint arXiv:2406.08481*, 2024a.

Yingyan Li, Shuyao Shang, Weisong Liu, Bing Zhan, Haochen Wang, Yuqi Wang, Yuntao Chen, Xiaoman Wang, Yasong An, Chufeng Tang, et al. Drivevla-w0: World models amplify data scaling law in autonomous driving. *arXiv preprint arXiv:2510.12796*, 2025d.

Yingyan Li, Yuqi Wang, Yang Liu, Jiawei He, Lue Fan, and Zhaoxiang Zhang. End-to-end driving with online trajectory evaluation via bev world model. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 27137–27146, 2025e.

Yongkang Li, Kaixin Xiong, Xiangyu Guo, Fang Li, Sixu Yan, Gangwei Xu, Lijun Zhou, Long Chen, Haiyang Sun, Bing Wang, et al. Recogdrive: A reinforced cognitive framework for end-to-end autonomous driving. *arXiv preprint arXiv:2506.08052*, 2025f.

Zhenxin Li, Kailin Li, Shihao Wang, Shiyi Lan, Zhiding Yu, Yishen Ji, Zhiqi Li, Ziyue Zhu, Jan Kautz, Zuxuan Wu, et al. Hydra-mdp: End-to-end multimodal planning with multi-target hydra-distillation. *arXiv preprint arXiv:2406.06978*, 2024b.

Zhenxin Li, Shihao Wang, Shiyi Lan, Zhiding Yu, Zuxuan Wu, and Jose M Alvarez. Hydra-next: Robust closed-loop driving with open-loop training. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 27305–27314, 2025g.

Zhenxin Li, Wenhao Yao, Zi Wang, Xinglong Sun, Joshua Chen, Nadine Chang, Maying Shen, Zuxuan Wu, Shiyi Lan, and Jose M Alvarez. Generalized trajectory scoring for end-to-end multimodal planning. *arXiv preprint arXiv:2506.06664*, 2025h.

Zhiqi Li, Zhiding Yu, Shiyi Lan, Jiahan Li, Jan Kautz, Tong Lu, and Jose M Alvarez. Is ego status all you need for open-loop end-to-end autonomous driving? In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 14864–14873, 2024c.

Bencheng Liao, Shaoyu Chen, Haoran Yin, Bo Jiang, Cheng Wang, Sixu Yan, Xinbang Zhang, Xiangyu Li, Ying Zhang, Qian Zhang, et al. Diffusiondrive: Truncated diffusion model for end-to-end autonomous driving. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pp. 12037–12047, 2025.

Lin Liu, Caiyan Jia, Guanyi Yu, Ziying Song, JunQiao Li, Feiyang Jia, Peiliang Wu, Xiaoshuai Hao, and Yadan Luo. Guideflow: Constraint-guided flow matching for planning in end-to-end autonomous driving. *arXiv preprint arXiv:2511.18729*, 2025a.

Shu Liu, Wenlin Chen, Weihao Li, Zheng Wang, Lijin Yang, Jianing Huang, Yipin Zhang, Zhongzhan Huang, Ze Cheng, and Hao Yang. Bridgedrive: Diffusion bridge policy for closed-loop trajectory planning in autonomous driving. *arXiv preprint arXiv:2509.23589*, 2025b.

Yuechen Luo, Fang Li, Shaoqing Xu, Zhiyi Lai, Lei Yang, Qimao Chen, Ziang Luo, Zixun Xie, Shengyin Jiang, Jiaxin Liu, et al. Adathinkdrive: Adaptive thinking via reinforcement learning for autonomous driving. *arXiv preprint arXiv:2509.13769*, 2025.

Long Nguyen, Micha Fauth, Bernhard Jaeger, Daniel Dauner, Maximilian Igl, Andreas Geiger, and Kashyap Chitta. Lead: Minimizing learner-expert asymmetry in end-to-end driving. *arXiv preprint arXiv:2512.20563*, 2025.

Egbert Otten. Inverse and forward dynamics: models of multi–body systems. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 358(1437):1493–1500, 2003.

Aditya Prakash, Kashyap Chitta, and Andreas Geiger. Multi-modal fusion transformer for end-to-end autonomous driving. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 7077–7087, 2021.

Ishaan Rawal, Shubh Gupta, Yihan Hu, and Wei Zhan. Nord: A data-efficient vision-language-action model that drives without reasoning. *arXiv preprint arXiv:2602.21172*, 2026.

Shuyao Shang, Yuntao Chen, Yuqi Wang, Yingyan Li, and ZHAO-XIANG ZHANG. Drivedpo: Policy learning via safety dpo for end-to-end autonomous driving. *Advances in Neural Information Processing Systems*, 38:81565–81585, 2026.

Matthew Strong, Wei-Jer Chang, Quentin Herau, Jiezhi Yang, Yihan Hu, Chensheng Peng, and Wei Zhan. Learning to drive is a free gift: Large-scale label-free autonomy pretraining from unposed in-the-wild videos. *arXiv preprint arXiv:2602.22091*, 2026.

Haisheng Su, Wei Wu, Feixiang Song, Junjie Zhang, Zhenjie Yang, and Junchi Yan. Drivemamba: Task-centric scalable state space model for efficient end-to-end autonomous driving. *arXiv preprint arXiv:2602.13301*, 2026.

Bin Sun, Yaoguang Cao, Yan Wang, Rui Wang, Jiachen Shang, Xiejie Feng, Jiayi Lu, Jia Shi, Shichun Yang, Xiaoyu Yan, et al. Minddrive: An all-in-one framework bridging world models and vision-language model for end-to-end autonomous driving. *arXiv preprint arXiv:2512.04441*, 2025a.

Wenchao Sun, Xuewu Lin, Yining Shi, Chuang Zhang, Haoran Wu, and Sifa Zheng. Sparsedrive: End-to-end autonomous driving via sparse scene representation. In *2025 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 8795–8801. IEEE, 2025b.

Wenchao Sun, Xuewu Lin, Keyu Chen, Zixiang Pei, Xiang Li, Yining Shi, and Sifa Zheng. Sparsedrivev2: Scoring is all you need for end-to-end autonomous driving. In *European Conference on Computer Vision*, pp. 446–463. Springer, 2026a.

Zhigang Sun, Yiru Wang, Anqing Jiang, Shuo Wang, Yu Gao, Yuwen Heng, Shouyi Zhang, An He, Hao Jiang, Jinhao Chai, et al. Diffsemanticfusion: Semantic raster bev fusion for autonomous driving via online map diffusion. *IEEE Robotics and Automation Letters*, 11(3):2354–2361, 2026b.

Jiacheng Tang, Zhiyuan Zhou, Zhuolin He, Jia Zhang, Kai Zhang, and Jian Pu. Causal-vad: De-confounding end-to-end autonomous driving via causal intervention. *arXiv preprint arXiv:2603.18561*, 2026.

Xiaolong Tang, Meina Kan, Shiguang Shan, and Xilin Chen. Plan-r1: Safe and feasible trajectory planning as language modeling. *arXiv preprint arXiv:2505.17659*, 2025.

Haochen Tian, Tianyu Li, Haochen Liu, Jiazhi Yang, Yihang Qiu, Guang Li, Junli Wang, Yinfeng Gao, Zhang Zhang, Liang Wang, et al. Simscale: Learning to drive via real-world simulation at scale. *arXiv preprint arXiv:2511.23369*, 2025.

Yang Tian, Sizhe Yang, Jia Zeng, Ping Wang, Dahua Lin, Hao Dong, and Jiangmiao Pang. Predictive inverse dynamics models are scalable learners for robotic manipulation. *arXiv preprint arXiv:2412.15109*, 2024.

Jie Wang, Guang Li, Zhijian Huang, Chenxu Dang, Hangjun Ye, Yahong Han, and Long Chen. Vggdrive: Empowering vision-language models with cross-view geometric grounding for autonomous driving. *arXiv preprint arXiv:2602.20794*, 2026a.

Junli Wang, Yinan Zheng, Xueyi Liu, Zebin Xing, Pengfei Li, Guang Li, Kun Ma, Guang Chen, Hangjun Ye, Zhongpu Xia, et al. Meanfuser: Fast one-step multi-modal trajectory generation and adaptive reconstruction via meanflow for end-to-end autonomous driving. *arXiv preprint arXiv:2602.20060*, 2026b.

Shihao Wang, Yingfei Liu, Tiancai Wang, Ying Li, and Xiangyu Zhang. Exploring object-centric temporal modeling for efficient multi-view 3d object detection. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 3621–3631, 2023.

Xiaofeng Wang, Zheng Zhu, Guan Huang, Xinze Chen, Jiagang Zhu, and Jiwen Lu. Drivedreamer: Towards real-world-drive world models for autonomous driving. In *European conference on computer vision*, pp. 55–72. Springer, 2024a.

Yuqi Wang, Jiawei He, Lue Fan, Hongxin Li, Yuntao Chen, and Zhaoxiang Zhang. Driving into the future: Multiview visual forecasting and planning with world model for autonomous driving. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 14749–14759, 2024b.

Xinshuo Weng, Boris Ivanovic, Yan Wang, Yue Wang, and Marco Pavone. Para-drive: Parallelized architecture for real-time autonomous driving. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 15449–15458, 2024.

Maciej K Wozniak, Lianhang Liu, Yixi Cai, and Patric Jensfelt. Prix: Learning to plan from raw pixels for end-to-end autonomous driving. *arXiv preprint arXiv:2507.17596*, 2025.

Tianze Xia, Yongkang Li, Lijun Zhou, Jingfeng Yao, Kaixin Xiong, Haiyang Sun, Bing Wang, Kun Ma, Guang Chen, Hangjun Ye, et al. Drivelaw: Unifying planning and video generation in a latent driving world. *arXiv preprint arXiv:2512.23421*, 2025.

Zebin Xing, Xingyu Zhang, Yang Hu, Bo Jiang, Tong He, Qian Zhang, Xiaoxiao Long, and Wei Yin. Goalflow: Goal-driven flow matching for multimodal trajectories generation in end-to-end autonomous driving. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pp. 1602–1611, 2025.

Yifang Xu, Jiahao Cui, Feipeng Cai, Zhihao Zhu, Hanlin Shang, Shan Luan, Mingwang Xu, Neng Zhang, Yaoyi Li, Jia Cai, et al. Wam-flow: Parallel coarse-to-fine motion planning via discrete flow matching for autonomous driving. *arXiv preprint arXiv:2512.06112*, 2025.

Jiazhi Yang, Kashyap Chitta, Shenyuan Gao, Long Chen, Yuqian Shao, Xiaosong Jia, Hongyang Li, Andreas Geiger, Xiangyu Yue, and Li Chen. Resim: Reliable world simulation for autonomous driving. *arXiv preprint arXiv:2506.09981*, 2025.

Pengxuan Yang, Ben Lu, Zhongpu Xia, Chao Han, Yinfeng Gao, Teng Zhang, Kun Zhan, XianPeng Lang, Yupeng Zheng, and Qichao Zhang. Worldrft: Latent world model planning with reinforcement fine-tuning for autonomous driving. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 11649–11657, 2026.

Wenhao Yao, Zhenxin Li, Shiyi Lan, Zi Wang, Xinglong Sun, Jose M Alvarez, and Zuxuan Wu. Drivesuprim: Towards precise trajectory selection for end-to-end planning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 11910–11918, 2026.

Liuhan Yin, Runkun Ju, Guodong Guo, and Erkang Cheng. Diffrefiner: Coarse to fine trajectory planning via diffusion refinement with semantic interaction for end to end autonomous driving. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 12009–12017, 2026.

Zhenlong Yuan, Chengxuan Qian, Jing Tang, Rui Chen, Zijian Song, Lei Sun, Xiangxiang Chu, Yujun Cai, Dapeng Zhang, and Shuo Li. Autodrive-r^2: Incentivizing reasoning and self-reflection capacity for vla model in autonomous driving. In *International Conference on Learning Representations*, volume 2026, pp. 97934–97955, 2026.

Shuang Zeng, Xinyuan Chang, Mengwei Xie, Xinran Liu, Yifan Bai, Zheng Pan, Mu Xu, Xing Wei, and Ning Guo. Futuresightdrive: Thinking visually with spatio-temporal cot for autonomous driving. *arXiv preprint arXiv:2505.17685*, 2025.

Bozhou Zhang, Nan Song, Jingyu Li, Xiatian Zhu, Jiankang Deng, and Li Zhang. Future-aware end-to-end driving: Bidirectional modeling of trajectory planning and scene evolution. *arXiv preprint arXiv:2510.11092*, 2025a.

Jinqing Zhang, Zehua Fu, Zelin Xu, Wenying Dai, Qingjie Liu, and Yunhong Wang. Resworld: Temporal residual world model for end-to-end autonomous driving. *arXiv preprint arXiv:2602.10884*, 2026.

Kaiwen Zhang, Zhenyu Tang, Xiaotao Hu, Xingang Pan, Xiaoyang Guo, Yuan Liu, Jingwei Huang, Li Yuan, Qian Zhang, Xiao-Xiao Long, et al. Epona: Autoregressive diffusion world model for autonomous driving. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 27220–27230, 2025b.

Rui Zhao, Yuze Fan, Ziguo Chen, Fei Gao, and Zhenhai Gao. Diffe2e: Rethinking end-to-end driving with a hybrid diffusion-regression-classification policy. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025.

Weicheng Zheng, Xiaofei Mao, Nanfei Ye, Pengxiang Li, Kun Zhan, Xianpeng Lang, and Hang Zhao. Driveagent-r1: Advancing vlm-based autonomous driving with active perception and hybrid thinking. *arXiv preprint arXiv:2507.20879*, 2025a.

Yupeng Zheng, Pengxuan Yang, Zebin Xing, Qichao Zhang, Yuhang Zheng, Yinfeng Gao, Pengfei Li, Teng Zhang, Zhongpu Xia, Peng Jia, et al. World4drive: End-to-end autonomous driving via intention-aware physical latent world model. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 28632–28642, 2025b.

Jialv Zou, Shaoyu Chen, Bencheng Liao, Zhiyu Zheng, Yuehao Song, Lefei Zhang, Qian Zhang, Wenyu Liu, and Xinggang Wang. Diffusiondrivev2: Reinforcement learning-constrained truncated diffusion modeling in end-to-end autonomous driving. *arXiv preprint arXiv:2512.07745*, 2025.

A NOTATION

Table 6 summarizes the main notation used in IDOL. We use k to index trajectory candidates, s to index trajectory poses, u to index latent rollout steps, and ℓ to index query-refinement iterations. The final refinement iteration is $\ell = L - 1$.

Table 6: Main notation used in IDOL.

Symbol	Description
$K,\ S$	Number of trajectory anchors and future poses per trajectory.
$U,\ L$	Number of latent rollout steps and query-refinement iterations.
B	Current latent BEV representation.
$\mathcal{A} = \{\tau_k\}_{k=1}^{K}$	Trajectory-anchor vocabulary, with $\tau_k \in \mathbb{R}^{S\times 3}$ containing planar positions and headings.
τ^{GT}	Expert future trajectory in the current ego frame.
$a_k,\ z$	Encoded anchor feature and ego-state feature.
$e_{k,\mathrm{init}}$	Initial planning query conditioned on the ego state and anchor k.
$e_{k,\mathrm{in}}^{(\ell)},\ e_{k,\mathrm{ref}}^{(\ell)}$	Planning query before and after refinement at iteration ℓ.
$q_{k,\ell}^{(u)}$	Ego query at step u of the candidate-conditioned latent rollout.
$B_{k,\ell,\mathrm{cond}}^{(u)}$	Latent BEV state conditioned through ego-query injection.
$S_k^{(\ell)},\ g_k^{(\ell)}$	Spatial and global IDM features aggregated over rollout transitions.
$k^\star$	Index of the anchor closest to τ^{GT} under the winner-take-all assignment.
$E_{\mathrm{mot}}(\tau)$	Fixed, normalized motion representation of trajectory τ.
h_{mot}	Training-only head predicting a motion-feature residual from the final global IDM feature.
$\mathrm{sg}[\cdot]$	Stop-gradient operator applied to the supervision target.

B IMPLEMENTATION DETAILS

Hyperparameters. Table 7 summarizes the key hyperparameters used in IDOL. The trajectory-anchor vocabulary is constructed offline and kept fixed during training and evaluation. For ablation studies, all settings are unchanged unless the corresponding component is explicitly modified.

Table 7: Hyperparameters of IDOL.

Parameter	Value
Epochs	30
Image input size	256×1024
BEV latent resolution	8×8
Number of BEV queries	64
Latent feature dimension	256
Planning horizon	4 s
Waypoint interval	0.5 s
Number of future waypoints	8
Trajectory-anchor vocabulary size	256
Closed-loop refinement iterations	1
Learning rate	2×10^{-4}
Weight decay	1×10^{-4}

Inverse dynamics model. IDOL adopts a lightweight MLP-based inverse dynamics model to convert adjacent imagined BEV states into motion-aware transition cues. Given two consecutive latent BEV states, the model first forms a transition representation by concatenating their query features along the channel dimension. A three-stage MLP transition encoder is then applied: an initial fusion stage compresses the paired-state representation into the latent feature space, followed by residual

MLP transformations that refine query-wise transition patterns while preserving the original spatial layout. The resulting query-wise representation can be viewed as a *spatial dynamics map*, which captures local motion-relevant changes across BEV queries. IDOL further aggregates this spatial dynamics map into a compact *global dynamics feature*, which serves as a trajectory-level motion condition for modulating the anchor-conditioned ego planning feature through an MLN layer before trajectory offset prediction. This design keeps the inverse dynamics branch lightweight while allowing the planner to exploit both local transition patterns and global motion context from adjacent imagined futures.

Loss configuration. IDOL is trained with trajectory offset regression, reward supervision, BEV semantic supervision, and IDM motion-feature correction. Table 8 lists the loss coefficients, which remain fixed throughout training.

Table 8: Training losses and their coefficients.

Loss term	Role	Coefficient
$\mathcal{L}_{\text{off}}$	Trajectory offset regression	1.0
$\mathcal{L}_{\text{off-im}}$	Auxiliary imitation-reward supervision	0.1
$\mathcal{L}_{\text{im}}$	Main imitation-reward supervision	1.0
$\mathcal{L}_{\text{sim}}$	Simulator-metric reward supervision	1.0
$\mathcal{L}_{\text{map}}^{\text{cur}}$	Current BEV semantic supervision	10.0
$\mathcal{L}_{\text{map}}^{\text{fut}}$	Future BEV semantic supervision	0.1
$\mathcal{L}_{\text{corr}}$	IDM motion-feature correction	0.2

The offset loss $\mathcal{L}_{\text{off}}$ applies L1 regression to the trajectory correction of the anchor closest to the expert trajectory under winner-take-all assignment. The imitation losses $\mathcal{L}_{\text{off-im}}$ and $\mathcal{L}_{\text{im}}$ supervise the auxiliary and main reward heads using a soft target distribution derived from anchor-to-expert distances. The simulator loss $\mathcal{L}_{\text{sim}}$ sums binary cross-entropy losses over five cached planning metrics. The current and future BEV losses supervise semantic scene representations using focal loss with $\alpha = 0.5$ and $\gamma = 2$.

Finally, downstream planning losses supervise the IDM indirectly, leaving the motion information encoded in its features underconstrained. We introduce $\mathcal{L}_{\text{corr}}$ to explicitly encourage the final global IDM feature to encode the motion adjustment from a candidate anchor toward the expert trajectory. For a trajectory τ, we define the fixed motion representation

$$E_{\text{mot}}(\tau) = [v_{1:S}/\sigma_v,\ a_{1:S-1}/\sigma_a,\ \omega_{1:S}/\sigma_\omega]\,, \tag{14}$$

where v_s is the signed displacement projected onto the midpoint heading and divided by Δt, $a_s = (v_{s+1} - v_s)/\Delta t$, and ω_s is the wrapped heading increment divided by Δt. We prepend the current ego pose and use $\Delta t = 0.5\,\text{s}$, $S = 8$, and normalization scales $(\sigma_v, \sigma_a, \sigma_\omega) = (10\,\text{m/s}, 3\,\text{m/s}^2, 0.5\,\text{rad/s})$.

Let $k^\star$ denote the anchor closest to the expert trajectory under the same winner-take-all assignment as offset regression. The correction loss is

$$\mathcal{L}_{\text{corr}} = \ell_{\text{SL1}}\Big(h_{\text{mot}}(g_{k^\star}^{(L-1)}), \text{sg}\big[E_{\text{mot}}(\tau^{\text{GT}}) - E_{\text{mot}}(\tau_{k^\star})\big]\Big)\,, \tag{15}$$

where h_{mot} is an auxiliary MLP head and sg denotes stop-gradient. The operator ℓ_{SL1} computes masked Smooth L1 loss with unit transition threshold, averaging valid entries within each motion component and then averaging the three components. Invalid targets are excluded, and heading-rate supervision is masked when either the expert or winning-anchor interval speed has magnitude below $0.5\,\text{m/s}$. Only the winning candidate is supervised. This fixed kinematic target provides interpretable supervision for the IDM representation used in query refinement. We use $\lambda_{\text{corr}} = 0.2$, and the auxiliary head is discarded at inference.

C Benchmark and metrics details

Benchmark. All models are trained on the NAVSIM training split. We evaluate IDOL on the NAVSIM benchmark. The *navtest* split is the standard test split used to evaluate closed-loop planning performance under common urban driving scenarios. The *navhard* split is a more challenging

Table 9: Performance comparison on the NAVSIM v2 navtest split using extended closed-loop metrics. The best and second-best results among learned methods are highlighted in bold and underlined, respectively. Superscripts indicate methods using additional training signals or settings beyond standard imitation learning: * denotes VLA/VLM-based methods, † denotes reinforcement learning, reward-guided training, or reinforcement fine-tuning, and ‡ denotes extra synthetic, pseudo-expert, or QA-style data.

Method	NC↑	DAC↑	DDC↑	TLC↑	EP↑	TTC↑	LK↑	HC↑	EC↑	EPDMS↑
Human Agent	100.0	100.0	99.8	100.0	87.4	100.0	100.0	98.1	90.1	90.3
Ego Status MLP (Li et al., 2024c)	93.1	77.9	92.7	99.6	86.0	91.5	89.4	**98.3**	85.4	64.0
TransFuser (Chitta et al., 2022)	96.9	89.9	97.8	99.7	87.1	95.4	92.7	**98.3**	87.2	76.7
Hydra-MDP++ (Li et al., 2025b)	97.2	97.5	99.4	99.6	83.1	96.5	94.4	98.2	70.9	81.4
DriveSuprim (Yao et al., 2026)	97.5	96.5	99.4	99.6	88.4	96.6	95.5	**98.3**	77.0	83.1
ReCogDrive*,† (Li et al., 2025f)	98.3	95.2	99.5	99.8	87.1	97.5	96.6	**98.3**	86.5	83.6
DiffusionDrive (Liao et al., 2025)	98.2	95.9	99.4	99.8	87.5	97.3	96.8	**98.3**	87.7	84.5
GTRS-Dense + SimScale‡,† (Tian et al., 2025)	98.4	98.8	99.4	**99.9**	87.9	98.1	96.4	97.6	58.8	84.6
World4Drive (Zheng et al., 2025b)	97.8	96.3	99.4	99.8	88.3	97.1	**97.7**	98.0	53.9	84.8
Epona (Zhang et al., 2025b)	97.1	95.7	99.3	99.7	88.6	96.3	97.0	98.0	67.8	85.1
DiffusionDriveV2† (Zou et al., 2025)	97.7	96.6	99.2	99.8	**88.9**	97.2	96.0	97.8	**91.0**	85.5
VADv2 (Chen et al., 2024b)	98.0	98.3	99.4	99.8	87.1	96.8	95.2	**98.3**	88.1	85.8
DriveVLA-W0* (Li et al., 2025d)	98.5	**99.1**	98.0	99.7	86.4	98.1	93.2	97.9	58.9	86.1
SGDrive*,‡ (Li et al., 2026)	98.6	94.3	99.5	**99.9**	86.0	97.9	96.1	**98.3**	85.9	86.2
DiffRefiner (Yin et al., 2026)	98.5	97.4	99.6	99.8	87.6	97.7	**97.7**	**98.3**	86.2	86.2
WorldRFT† (Yang et al., 2026)	97.8	96.5	99.5	99.8	88.5	97.0	97.4	98.1	69.1	86.7
SafeDrive (Kim et al., 2026)	**99.5**	99.0	**99.7**	**99.9**	88.6	**98.9**	97.5	98.2	81.9	87.5
MeanFuser (Wang et al., 2026b)	98.3	97.2	99.6	99.8	87.6	97.4	97.3	**98.3**	88.2	89.5
IDOL	98.8	97.6	99.5	99.8	87.1	98.3	96.3	**98.3**	85.5	**89.6**

subset designed to stress-test planners under difficult cases such as complex interactions, stricter route following, and long-tail driving conditions. For NAVSIM v2, navhard is evaluated under the pseudo-simulation protocol, including the stage-1-only setting and the two-stage setting. The former measures the quality of the initially generated trajectory, while the latter further evaluates planning after pseudo-simulation rollout and therefore reflects closed-loop robustness more directly. Since point-cloud inputs are unavailable in stage 2, we separately train a camera-based IDOL variant using the Latent TransFuser (LTF) formulation. Specifically, the LiDAR input is replaced by a learnable latent tensor initialized from an LTF checkpoint and jointly optimized with the IDOL network on the NAVSIM training split, while retaining the fusion architecture and downstream planning modules. The resulting IDOL-LTF checkpoint is used throughout the two-stage evaluation; the model does not switch modalities between stages or rely on dynamic modality-dropout training.

PDMS. For NAVSIM v1, we report the Predictive Driver Model Score (PDMS), which evaluates trajectory quality after non-reactive simulation. PDMS uses no-collision (NC) and drivable area compliance (DAC) as multiplicative safety constraints, and combines ego progress (EP), time-to-collision within bound (TTC), and comfort (C) as weighted quality terms:

$$\mathrm{PDMS} = \mathrm{NC} \cdot \mathrm{DAC} \cdot \frac{5\mathrm{EP} + 5\mathrm{TTC} + 2\mathrm{C}}{12}. \tag{16}$$

All component metrics are normalized to $[0, 1]$, and the final reported PDMS is averaged over all evaluated scenarios and expressed as a percentage.

EPDMS. For NAVSIM v2, we report the Extended Predictive Driver Model Score (EPDMS). Compared with PDMS, EPDMS adds driving direction compliance (DDC) and traffic light compliance (TLC) as multiplicative rule-compliance constraints, and further introduces lane keeping (LK), history comfort (HC), and extended comfort (EC) into the weighted quality term. Following the NAVSIM v2 metric definition, we write the filtered value of each metric as $\tilde{m}$, where human-trajectory violations are filtered to avoid false-positive penalties. EPDMS is computed as

$$\mathrm{EPDMS} = \tilde{\mathrm{NC}} \cdot \tilde{\mathrm{DAC}} \cdot \tilde{\mathrm{DDC}} \cdot \tilde{\mathrm{TLC}} \cdot \frac{5\tilde{\mathrm{EP}} + 5\tilde{\mathrm{TTC}} + 2\tilde{\mathrm{LK}} + 2\tilde{\mathrm{HC}} + 2\tilde{\mathrm{EC}}}{16}. \tag{17}$$

EPDMS therefore extends PDMS from basic safety and progress evaluation to stricter rule compliance, lane keeping, and pseudo-closed-loop consistency, making it more suitable for NAVSIM v2 and the challenging navhard split.

D Additional evaluation on NAVSIM v2 and Bench2Drive

Evaluation on NAVSIM v2. We report extended NAVSIM v2 results on the standard navtest split and the more challenging navhard stage-1 setting. Additional baseline results are collected from publicly available papers (Jiang et al., 2025b; Sun et al., 2025a; 2026b; Liu et al., 2025a). As shown in Table 9, IDOL achieves 89.6 EPDMS on navtest, ranking first among learned methods and remaining competitive with the human agent, with consistently strong safety, rule-compliance, and comfort-related metrics.

Under the navhard stage-1 protocol in Table 10, IDOL achieves 76.2 EPDMS, outperforming prior learned planners including Hydra-MDP and WoTE. These results support the effectiveness of inverse-dynamics-guided future reasoning across both standard and challenging long-tail NAVSIM v2 settings.

Interactive closed-loop evaluation on Bench2Drive. To evaluate performance beyond NAVSIM, we additionally test IDOL on Bench2Drive, a CARLA-based benchmark comprising 220 routes and 44 safety-critical interactive scenarios. Following the official evaluation protocol, we report Driving Score (DS) and Success Rate (SR) to assess driving performance and task completion under continuous interaction with the environment.

As shown in Table 11, IDOL uses a ResNet-34 image backbone and achieves a DS of 63.22 and an SR of 34.09%, outperforming the listed planners, including the future-BEV-based methods SeerDrive and WoTE. Compared with WoTE, IDOL improves DS by 1.51 points and SR by 2.73 percentage points. These results provide further evidence that inverse-dynamics-guided future reasoning supports effective planning and task completion under interactive closed-loop evaluation beyond NAVSIM.

Table 10: Performance comparison on the NAVSIM v2 navhard split under the stage-1-only protocol. Bold and underlined values indicate the best and second-best results, respectively, for each metric. † indicates results reported with the V2-99 backbone.

Method	NC↑	DAC↑	DDC↑	TLC↑	EP↑	TTC↑	LK↑	HC↑	EC↑	EPDMS↑
TransFuser (Chitta et al., 2022)	96.3	74.6	98.4	99.3	82.9	93.7	92.7	97.5	<u>78.2</u>	60.5
LTF (Chitta et al., 2022)	96.2	79.5	<u>99.1</u>	99.5	<u>84.1</u>	95.1	94.2	97.5	**79.1**	62.3
DiffusionDrive (Liao et al., 2025)	95.9	84.0	98.6	**99.8**	**84.4**	96.0	<u>95.1</u>	<u>97.6</u>	71.1	63.2
WoTE (Li et al., 2025e)	<u>97.4</u>	88.2	97.8	99.3	82.7	96.4	90.9	97.3	68.0	66.7
Hydra-MDP† (Li et al., 2024b)	**97.6**	**96.4**	**99.2**	99.3	80.2	**96.9**	94.9	**97.8**	58.7	<u>73.1</u>
IDOL	97.2	<u>89.6</u>	98.0	<u>99.6</u>	82.3	**96.9**	**95.3**	<u>97.6</u>	72.9	**76.2**

Table 11: Interactive closed-loop evaluation on Bench2Drive. Higher Driving Score (DS) and Success Rate (SR) indicate better performance. Bold values denote the best results among the listed methods.

Method	Driving Score (DS) ↑	Success Rate (SR) ↑
VAD (Jiang et al., 2023)	42.35	15.00%
UniAD-Base (Hu et al., 2023b)	45.81	16.36%
SeerDrive (Zhang et al., 2025a)	58.32	30.17%
WoTE (Li et al., 2025e)	61.71	31.36%
IDOL	**63.22**	**34.09%**

E Additional ablation studies

We further examine whether learned inverse dynamics provides better planning guidance than simpler future-state representations, whether its gains arise from additional model capacity, how global

inverse dynamics features should be fused with the refined planning query, and how performance varies with the trajectory-anchor vocabulary size.

Ablation on transition-aware future modeling. We compare three representations for planning refinement: imagined future BEV states without explicit transition modeling, direct latent differences between adjacent future states, and IDM features learned from ordered pairs of adjacent states. To control for model capacity, we additionally apply parameter-matched MLPs to the future-state-only and latent-difference representations. Both augmented variants have exactly the same total parameter count as IDOL. Across these comparisons, the world model, spatial cross-attention, global modulation, iterative refinement, trajectory anchors, losses, and training settings remain unchanged.

As shown in Table 12, latent differences improve PDMS from 88.6 to 89.0 over future states alone, supporting the value of transition information. Adding parameter-matched MLPs yields only modest gains of 0.3 and 0.1 points, respectively. IDM achieves 90.0 PDMS, exceeding the two parameter-matched variants by 1.1 and 0.9 points. These results indicate that additional parameters alone do not explain the improvement. While a future-state MLP operates on individual states and latent subtraction supplies a fixed, coordinate-wise change signal, IDM jointly processes ordered state pairs, preserving endpoint context and learning nonlinear, state-conditioned transition cues for planning refinement.

Table 12: Ablation of transition-aware future modeling on the NAVSIM navtest split. Parameter-matched MLP variants have the same total parameter count as IDOL; all other components and training settings are fixed. Parameter counts are reported in millions and rounded to one decimal place.

Transition Modeling	**Params (M)**	**NC ↑**	**DAC ↑**	**TTC ↑**	**Comf. ↑**	**EP ↑**	**PDMS ↑**
Future State Only	68.2	98.4	96.6	95.2	**100**	82.6	88.6
Future State Only + Parameter-Matched MLP	69.4	98.5	97.0	94.7	**100**	83.3	88.9
Latent Difference	68.2	98.6	96.8	95.7	**100**	82.8	89.0
Latent Difference + Parameter-Matched MLP	69.4	**98.7**	97.1	**95.8**	**100**	82.6	89.1
IDM	69.4	**98.7**	**97.7**	95.5	**100**	**84.0**	**90.0**

Ablation on global dynamics fusion. We further study how to fuse the global dynamics feature produced by the inverse dynamics model. In the inverse-dynamics-guided refinement network, the global dynamics feature provides holistic calibration for the query refined by spatial transition evidence. We compare three simple fusion strategies: additive fusion, which directly injects the projected global feature into the query; concat-MLP fusion, which learns a nonlinear fusion from the concatenated query and global feature; and the MLN-based fusion used in IDOL. As shown in Table 13, all variants obtain strong performance, indicating that the global inverse-dynamics cue is consistently beneficial for query refinement. These results verify that IDOL's performance is not only due to using inverse-dynamics features, but also benefits from an effective query-calibration mechanism that better integrates global transition information into planning.

Table 13: Ablation of global dynamics fusion strategies in the inverse-dynamics-guided refinement network on the NAVSIM navtest split. All variants use the same spatial branch and only replace the fusion operation between the refined ego query and the global dynamics feature.

Fusion Strategy	**NC ↑**	**DAC ↑**	**TTC ↑**	**Comf. ↑**	**EP ↑**	**PDMS ↑**
Additive	**98.7**	97.2	**95.7**	**100**	83.5	89.5
Concat-MLP	**98.7**	97.6	95.5	**100**	83.8	89.8
MLN	**98.7**	**97.7**	95.5	**100**	**84.0**	**90.0**

Sensitivity to anchor-set size. We evaluate IDOL with $K = 128$, $K = 256$, and $K = 512$ trajectory anchors, keeping the anchor-construction procedure, backbone, world model, IDM, losses, and training settings unchanged. As shown in Table 14, PDMS ranges from 89.6 to 90.0, indicating limited sensitivity to anchor-set size within the tested range. A smaller vocabulary may limit trajectory coverage, whereas a denser vocabulary may introduce redundant candidates and increase

ranking difficulty. We retain $K = 256$ as the default, as it achieves the highest PDMS while using fewer candidates than $K = 512$.

Table 14: Sensitivity to trajectory-anchor vocabulary size on the NAVSIM navtest split. Only the number of anchors is varied; the anchor-construction procedure and all other model and training settings remain unchanged.

Number of Anchors	NC ↑	DAC ↑	TTC ↑	Comf. ↑	EP ↑	PDMS ↑
128	**98.7**	97.6	**95.5**	**100**	83.8	89.8
256	**98.7**	**97.7**	**95.5**	**100**	**84.0**	**90.0**
512	**98.7**	97.2	95.3	**100**	**84.0**	89.6

F ADDITIONAL QUALITATIVE VISUALIZATION

To further demonstrate the effectiveness of our proposed method, we expand the qualitative visualization experiments with additional representative driving scenarios. Figure 4 presents representative visualization results on the NAVSIM navtest split under three driving maneuvers, including straight driving, left turn, and right turn. Each example contains four panels that illustrate how the proposed IDM module extracts spatial dynamics from latent BEV transitions and uses them to refine the trajectory prediction.

Panel (a) shows the latent BEV change between two adjacent imagined BEV states produced by the world model. The highlighted regions indicate where the scene representation changes most significantly from frame t to frame $t + 1$. Therefore, this panel serves as the reference evidence for spatially localized future dynamics, such as ego-motion-induced changes, nearby traffic participants, and lane-level scene evolution.

Panel (b) visualizes the spatial dynamics inferred by the IDM from the same adjacent BEV pair. Instead of directly displaying a semantic map, this panel shows the inverse-dynamics response over BEV queries. The white contour marks the high-change regions extracted from Panel (a), and is overlaid as a reference for comparing IDM responses with latent BEV transitions. The regions with stronger activation correspond to locations where IDM considers the transition between the two BEV states informative for explaining the underlying motion. The visual consistency between Panel (a) and Panel (b) suggests that IDM is not responding to arbitrary static background features, but is sensitive to the spatial regions where the latent scene actually evolves.

Panel (c) further shows the attention distribution from the trajectory query to the IDM spatial dynamics. While Panel (b) indicates where inverse-dynamics information is extracted, Panel (c) reveals which dynamic regions are actually attended to by the refinement query. The concentration of attention around the changing BEV regions demonstrates that the refinement process actively uses the IDM-derived spatial evidence, rather than treating IDM as an isolated auxiliary branch.

Panel (d) compares the original trajectory, the expert trajectory, and the refined trajectory. The original trajectory is decoded before applying the IDM-based closed-loop refinement, while the refined trajectory is decoded after the query has absorbed inverse-dynamics information through the IDM attention and fusion modules. Across straight, left-turn, and right-turn cases, the refined trajectory is visually better aligned with the expert trajectory, showing that the IDM-guided update helps correct the initial prediction toward a more plausible future motion.

Overall, these qualitative results provide direct visual evidence for the effectiveness of the IDM module. The latent BEV transition identifies where the future scene changes, the IDM spatial response recovers dynamics from those changing regions, and the query attention confirms that such information is used by the planner. The resulting trajectory refinement illustrates the practical value of this mechanism: IDM helps the model focus on motion-relevant spatial changes and improves the final trajectory prediction in different driving maneuvers. This supports our hypothesis that inverse-dynamics reasoning provides a meaningful bridge between latent world-model prediction and trajectory-level decision refinement.

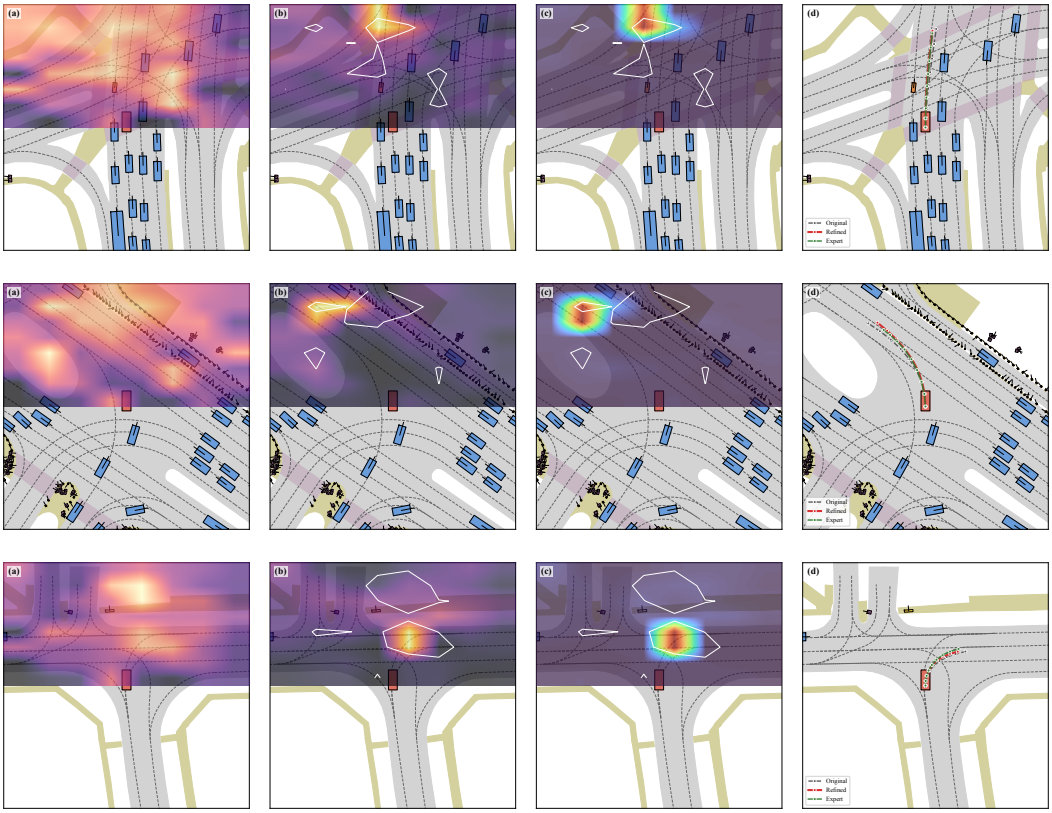


Figure 4: Qualitative visualization of inverse-dynamics refinement under three representative driving maneuvers on the NAVSIM navtest split. From top to bottom: straight driving, left turn, and right turn. In each row, panels (a)–(d) show the latent BEV change, the IDM dynamics inferred from adjacent BEV states, the IDM attention on dynamic queries, and the resulting refined trajectory, respectively.

G Failure Cases

We further analyze two representative failure cases on navtest, as shown in Figure 5. Each row contains the front-view camera image and the corresponding BEV trajectory comparison. The refined trajectory denotes the final prediction after IDM-based refinement, and the expert trajectory is used as the reference.

In the first case, the ego vehicle approaches a turning area with ambiguous road geometry. The refined trajectory deviates clearly from the expert trajectory and follows an incorrect spatial direction. This indicates that when the model fails to infer the correct maneuver intention, IDM refinement alone may not fully recover the desired path.

In the second case, the scene involves a complex intersection or roundabout-like structure. Although the refined trajectory captures part of the motion trend, it remains spatially offset from the expert trajectory. This suggests that topology-sensitive scenarios still require more precise reasoning about lane connectivity, road curvature, and maneuver timing.

These cases show that IDM provides useful dynamic cues, but its effectiveness depends on whether such cues can be correctly fused with scene topology and trajectory intent. Future work may improve these limitations with stronger topology-aware constraints and uncertainty-aware refinement.



Figure 5: Representative failure cases on navtest.

H LIMITATIONS AND FUTURE WORK

Although IDOL shows strong performance across NAVSIM benchmarks, several limitations remain.

Model capacity. Our inverse dynamics model is designed to be lightweight, which keeps the framework efficient but may limit its ability to capture more complex long-horizon interactions and subtle motion patterns. Future work could scale the IDM with larger temporal contexts, stronger transition modeling, or pretraining on broader driving sequences.

Dependence on latent future prediction. IDOL relies on the quality of the imagined latent BEV futures produced by the world model. When future prediction is inaccurate in highly ambiguous or rare scenarios, the decoded inverse-dynamics cues may also become less reliable. Improving the robustness of latent world modeling is therefore an important direction.

Fixed refinement strategy. The current closed-loop refinement uses a fixed refinement setting, while different scenes may require different levels of future-aware reasoning. Adaptive refinement strategies could further improve stability and avoid unnecessary iterative updates.

Broader evaluation. Although our experiments cover both standard and challenging NAVSIM settings, future work will further evaluate inverse-dynamics-guided planning in more interactive, long-tail, and real-world driving scenarios.